A Bayesian approach to sapwood estimates and felling dates in dendrochronology

Andrew Millard

Department of Archaeology, University of Durham, South Road, Durham. DH1 3LE. UK

Abstract

An improved method of generating sapwood estimates for oak is developed. This suggests a revision of the 95% confidence range from 10-40 to 9-36 rings for trees from Southern England. Current methods for estimating felling dates on timbers with incomplete sapwood do not generate true 95% confidence limits, and a Bayesian method for deriving such limits is presented. For timbers with no sapwood the addition of 12 years to the date of the final ring is shown to give a 95% confidence limit on the *terminus post quem* for felling. The further application of these methods is illustrated by calculation of the common felling date for timbers from the Great Kitchen at Windsor Castle.

Introduction

In dendrochronology dates are produced by matching ring-width sequences of timbers to master chronologies to give the year of growth of each ring. The method has the potential to give extremely precise dates for archaeological events, if suitable timbers are available with the bark on them. Frequently, however, the bark is not present and a felling date has to be estimated as a 95% confidence interval using the date of the heartwood-sapwood boundary and an estimate for the number of sapwood rings in a tree. (Baillie 1982; Hillam 1998).

Various sapwood estimates have been suggested on the basis of studies of modern trees (Hughes *et al.* 1981), and archaeological samples with complete sapwood (Hillam *et al.* 1987;

Miles 1997; Hillam 1998, 11). The distribution of number of sapwood rings in these samples of trees has been found to conform to a log-normal distribution (Hillam *et al.* 1987), and a 95% confidence interval on the log values is used to give the sapwood estimate. Hence the standard procedure is that if the date of the heartwood-sapwood boundary on a timber can be established but little or no sapwood survives, then the felling date is estimated as 10-55 years later (using the Sheffield sapwood estimate of Hillam *et al.* (1987), though others are used elsewhere - see the review of Miles (1997)). If a number of sapwood rings survive, *e.g.* 25, the felling date is estimated a between then and the end of the standard sapwood estimate, in this case 25-55 years after the heartwood-sapwood boundary (Laxton and Litton 1988, 53-54; Hillam 1998, 27). If the heartwood-sapwood boundary is missing then the lower limit of the sapwood estimate is added to the last ring to give a *terminus post quem* for felling.

This paper seeks to view these estimates from a Bayesian perspective (Buck *et al.* 1996), revising the normal sapwood estimate, suggesting how the extra available data of the observed number of sapwood rings should be incorporated into the date and suggesting a revised method for calculating a *terminus post quem* for felling. By generating probability density functions for dates they can be incorporated into Bayesian analyses, and this is illustrated by the development of a more rigorous analysis of the felling date of a group of timbers which are assumed to have been felled contemporaneously.

Simple sapwood estimates

Sapwood estimates are usually derived by the method of Hughes *et al.* (1981) which calculates a 95% confidence interval on the normally-distributed log-transformed sapwood ring count data for a group of trees and then converts this back into a confidence interval expressed in numbers of rings, which are finally rounded to the nearest integer to give the working values. If the number of sapwood rings was a continuous variable this would be satisfactory, but in fact it is a discrete variable. Having made a log-normal approximation to the distribution, the best estimate of the probability of having *n* rings is the integral of the log-normal distribution from *n*-½ to *n*+½ and these values for discrete numbers of rings should be used in subsequent evaluation of confidence intervals.

For the purposes of illustrating the method the data of Miles (1997) for the numbers of sapwood rings in 406 medieval and post-medieval oak timbers from Southern England is used here. Thus the results only apply when this collection is the appropriate comparator for the particular timber(s) under consideration. However the method can be transferred to other datasets relatively easily, so that other workers can produce sapwood estimates for the regions that interest them.

Figure 1 shows the data from Miles together with the probabilities calculated from a normal distribution with the same mean and standard deviation as the log-transformed data. The "95% mid-range" calculated by the method of Hughes *et al.* (1981) is 10-40 rings. It is clear from the figure that 9 rings are more likely than 37, 38, 39 or 40 rings. The preferred 95% confidence interval from a Bayesian perspective is the highest probability density (hpd) region, which in this case is 9-36 rings (95.2% confidence). Working with a discrete probability function means that it is not always possible to calculate a hpd region with exactly 95% confidence, so in this paper, the smallest hpd with at least 95% confidence is calculated and the actual confidence indicated in parenthesis.

Sapwood estimates with surviving sapwood

The standard method of simply truncating the sapwood estimate at the observed number of rings does not yield an interval with a fixed confidence level. Hillam (1998, 27) states that the truncated "felling date range" is a 95% confidence estimate, but this is not necessarily the case. Such truncated date ranges represent confidence limits whose level depends on the observed number of rings. Thus it is not possible to compare two felling date ranges on the basis that they have the same confidence level. A Bayesian approach to this problem may be developed as follows. If we wish to estimate the number of sapwood rings then *a priori* the probability density is given by the log-normal distribution as discussed above, *i.e.*,

$$p(n|\lambda,\delta) = \int_{n-½}^{n+½} \frac{1}{x\delta\sqrt{2\pi}} \exp\left(-\frac{(\ln x-\lambda)^2}{2\delta}\right) \quad (1)$$

where λ and δ are the parameters of the distribution. In addition we know that there are *a* rings present, and thus the likelihood is given by

$$l(a;n) \propto \begin{cases} 1 \text{ for } a \leq n \\ 0 \text{ for } a < n \end{cases}$$

and thus the posterior probability is given by

$$p(n\,|\,a) \propto l(a;n)p(n\,|\,\lambda,\delta)$$

This distribution can be numerically normalised to give probabilities and hence 95% hpd ranges. This is illustrated in Figure 2, which shows the effect of a=25 on the distribution of Figure 1. The hpd region is now 25-43 rings (95.5%) rather than the 25-40 rings (92.5%) of the truncation method.

Estimating a *terminus post quem*

Currently the *terminus post quem* (*tpq)* for the felling date of a tree without a heartwood-sapwood boundary is estimated by adding the lower limit of the sapwood estimate to the date of the latest ring. This gives a *tpq* with a confidence of more than 95%, the precise amount depending on the parameters of the distribution of sapwood ring numbers. For example, with the distribution shown in Figure 1, adding 10 years (rings) gives a *tpq* with 98.6% confidence, adding the new, lower limit of 9 rings gives a *tpq* with 99.4% confidence. Adding 12 years gives a *tpq* with 94.9% confidence, and is thus recommended as the closest to a 95% confidence limit that can be achieved given the discrete nature of the distribution.

Obtaining felling date ranges

As the new method outlined in this paper gives probability density functions for felling dates based on sapwood estimates, these can be used to incorporate dendrochronological dates into a wide range of other Bayesian analyses (*e.g.* using the program OxCal (Bronk Ramsey 1995)). This is demonstrated here using a purely dendrochronological example: the estimation of the single year of felling of a group of timbers with and without sapwood.

Those timbers with no heartwood-sapwood boundary give a *tpq* for the felling. If H_i is the final observed heartwood ring in timber i, then the likelihood is given by[1]:

$$l(H_i;\theta) \propto \begin{cases} 1 & \text{for } H_i < \theta \\ 0 & \text{for } H_i \geq \theta \end{cases}$$

Those timbers with a heartwood-sapwood boundary, but no sapwood rings have a likelihood given by:

$$l(\theta\; H_i) \propto \begin{cases} p(\theta - H_i|\lambda, \delta) & \text{for } \theta > H_i \\ 0 & \text{for } \theta \leq H_i \end{cases}$$

with $p(n|\lambda, \delta)$ as in equation 1. For those timbers with a_i sapwood rings, the likelihood is given by:

$$l(H_i;\theta) \propto \begin{cases} p(\theta - H_i | \lambda, \delta) & \text{for } a_i \leq \theta - H_i \\ 0 & \text{for } a_i > \theta - H_i \end{cases}$$

Taking a vague prior (*i.e.* preferring no year above another initially), the combined posterior is given by:

$$p(\theta \mid \mathbf{H},\mathbf{a}) \propto \prod_i l(H_i, a_i; \theta)$$

The usefulness of this approach may be illustrated by an example. Table 1 summarises the date of the last heartwood ring and the extent of surviving sapwood for 14 timbers from the Great Kitchen at Windsor Castle (Hillam 1998). The closeness of the end years or heartwood-sapwood boundaries suggests that the timbers are all contemporary. The historical question is: What was the year of felling?

Figure 3 shows the probability distributions for the felling dates of those timbers with heartwood/sapwood boundaries and the combined posterior probability density. In this case the timbers with no sapwood do not contribute anything to the result and are therefore not shown. The estimated date of felling is 1574-1582 (97.5% confidence). Using the method of Hillam (1998, 11) and the sapwood estimate of 10-40 years derived from the method of Hughes *et al.* (1981) on Miles (1997) data would have resulted in a range of 1576-1588. It so happens that in this structure the true date is known because an additional timber with bark (12070B, not included in the dataset here) gives a felling date of 1577 AD and this is confirmed by documentary records of repairs in that year (Hillam 1998, 16). Both calculations give ranges which agree with the known date, but the Bayesian method is preferable because the results are a statistically well defined hpd region with specific confidence limits. The new method yields a 95% confidence range of 9 years compared to 13 years from the previous method. As Pearson (1997) observed with regard to a 5 year revision of the sapwood estimate for Kent "If one is going in for accurate dating, this matters".

However although the results are statistically more meaningful the quality of sapwood estimates and felling dates so obtained is still subject to caveats. Notable assumptions are (a) the appropriateness of the known sapwood sample in both its geographical and temporal range, and (b) the absence of a plantation effect, where the numbers of sapwood rings of different timbers are correlated due to similar growing conditions. Additionally, the use of a log-normal distribution is empirically based without any explicit basis in plant physiology. Further improvements in sapwood estimates will come from increased understanding of biological, geographical and temporal variation in sapwood ring numbers.

Conclusion

Bayesian statistical methods are finding increasing application in chronometric and archaeological studies (Litton and Buck 1995). This note has demonstrated how they may be applied to sapwood estimates in dendrochronology to provide statistically more rigorous descriptions of estimated felling dates for both individual timbers and groups of timbers. They have the advantage, in many cases, of also producing shorter ranges for estimated dates, which makes them attractive to the end user, who can know that the new methods are both statistically better and more precise. In addition the production ofprobability distributions now allows the dates to be integrated with other dating evidence (*e.g.* radiocarbon dates) within a Bayesian framework.

Note

1. This formula is a simplification. In principle one could use the information embodied in equation 1 to generate a more complex probability distribution but simple combination of

these leads to a bias to later dates which worsens as the number of *tpq*s increases. The solution to this problem is not clear at present. I am grateful to C Bronk Ramsey for pointing out this problem with multiple complex *tpq*.

References

Baillie, M. G. L., 1982, *Tree-ring dating and archaeology*, University of Chicago Press, London.

Bronk Ramsey, C., 1995, Radiocarbon calibration and stratigraphy: the OxCal program, *Radiocarbon*, **37**, 425-430

Buck, C. E., Cavanagh, W. G., and Litton, C. D., 1996, *Bayesian approach to interpreting archaeological data*, John Wiley, Chichester.

Hillam, J., 1998, *Dendrochronology: guidelines on producing and interpreting dendrochronological dates*, English Heritage, London.

Hillam, J., Morgan, R. A., and Tyers, I., 1987, Sapwood estimates and the dating of short ring sequences, in *Applications of tree-ring studies: current research in dendrochronology and related areas* (ed. R. G. W. Ward), British Archaeological Reports International Series, **S333**, 165-185

Hughes, M. K., Milsom, S. J., and Leggett, P. A., 1981, Sapwood estimates in the interpretation of tree-ring dates, *Journal of Archaeological Science*, **8**, 381-390

Laxton, R. R., and Litton, C. D., 1988, *An East Midlands master tree-ring chronology and its use for the dating of vernacular buildings*, University of Nottingham, Department of Classical and Archaeological Studies, Monograph Series, **III.**

Litton, C. D., and Buck, C. E., 1995, The Bayesian approach to the interpretation of archaeological data, *Archaeometry*, **37**, 1-24.

Miles, D., 1997, The interpretation, presentation and use of tree-ring dates, *Vernacular Architecture*, **28**, 40-56. (As amended by erratum slip in *Vernacular Architecture* **29**.)

Pearson, S., 1997, Tree-ring dating: a review, *Vernacular Architecture* **28**, 25-39.

Table 1

Summary of dendrochronological data of timbers from the Great Kitchen, Windsor Castle.

(Derived from Hillam (1998) Figure 12.)

Sample	last heartwood ring (AD)	heartwood-sapwood boundary?	surviving sapwood rings
12013	1535	N	
12055	1537	N	
12009	1548	N	
12054	1549	Y	0
12062	1550	N	
12057	1552	Y	2
12049	1552	Y	14
12003	1553	N	
12066	1553	N	
12067	1554	Y	0
12017	1557	N	
12060	1559	Y	12
12051	1565	N	
12024	1566	Y	7

Figure captions

Figure 1: Distribution of number of sapwood rings. Bars show frequency data of Miles (1997) for southern England, dashes show probabilities calculated as described in the text.

Figure 2: Posterior probability distribution for number of sapwood rings, given that 25 sapwood rings have been observed.

Figure 3: Probability distributions for the felling date of timbers from the Great Kitchen, Windsor Castle.

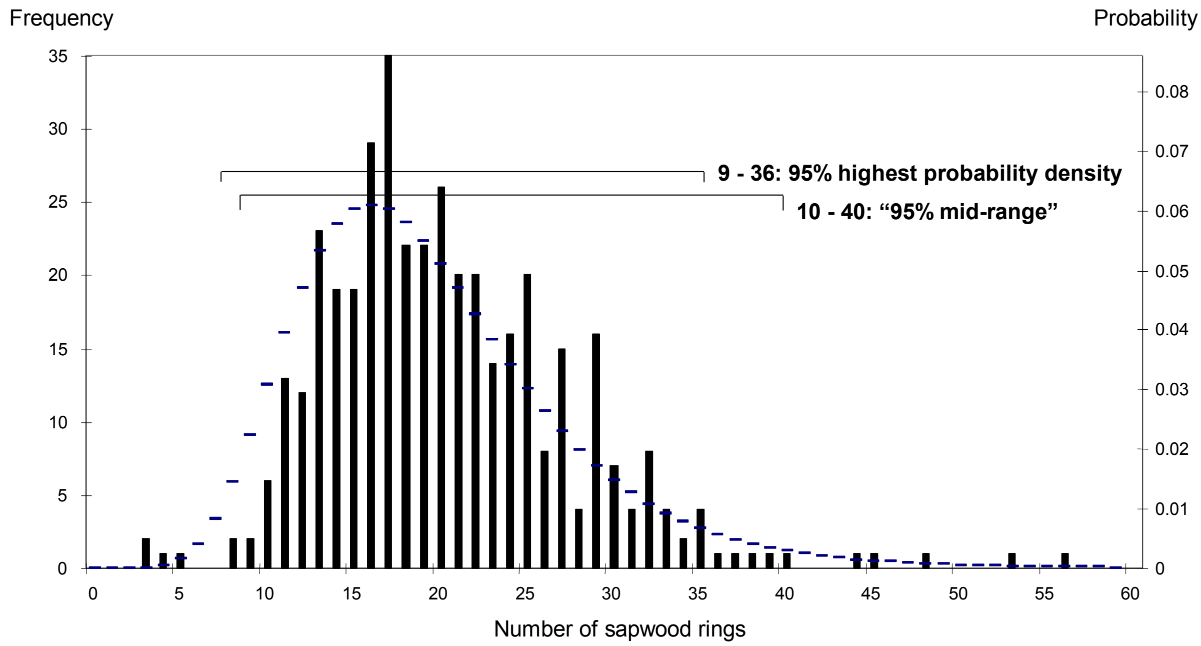

Figure 1: Distribution of number of sapwood rings. Bars show frequency data of Miles (1997) for southern England, dashes show probabilities calculated as described in the text.

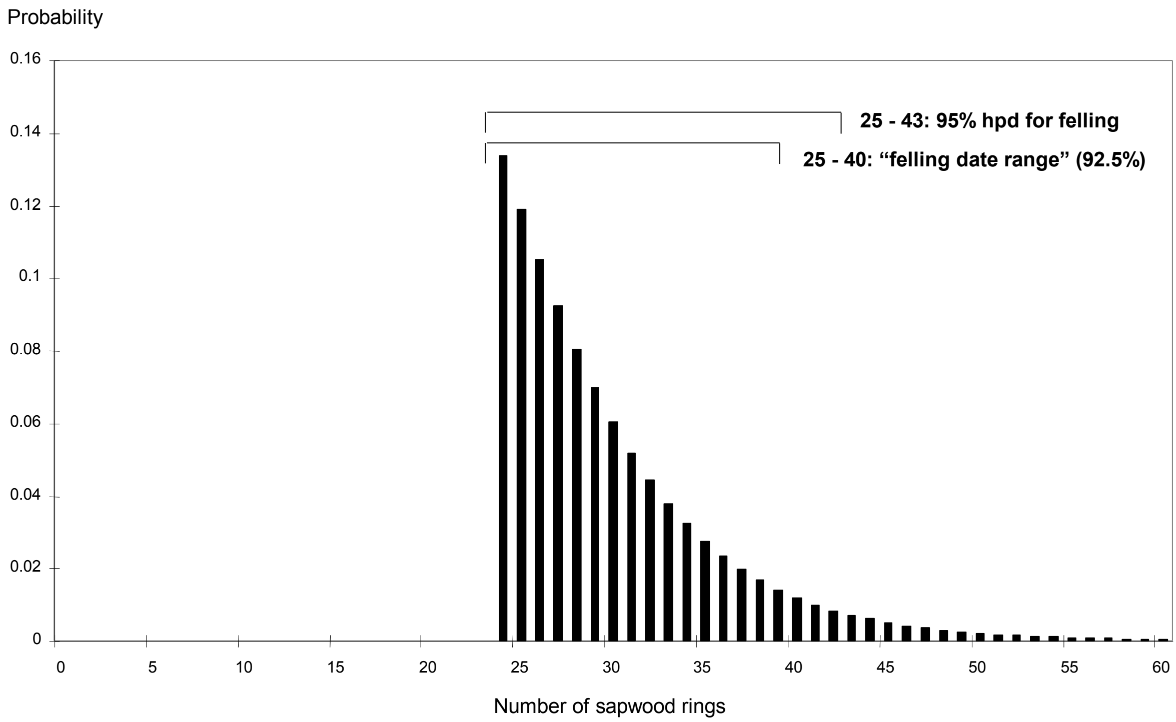

Figure 2: Posterior probability distribution for number of sapwood rings, given that 25 sapwood rings have been observed.

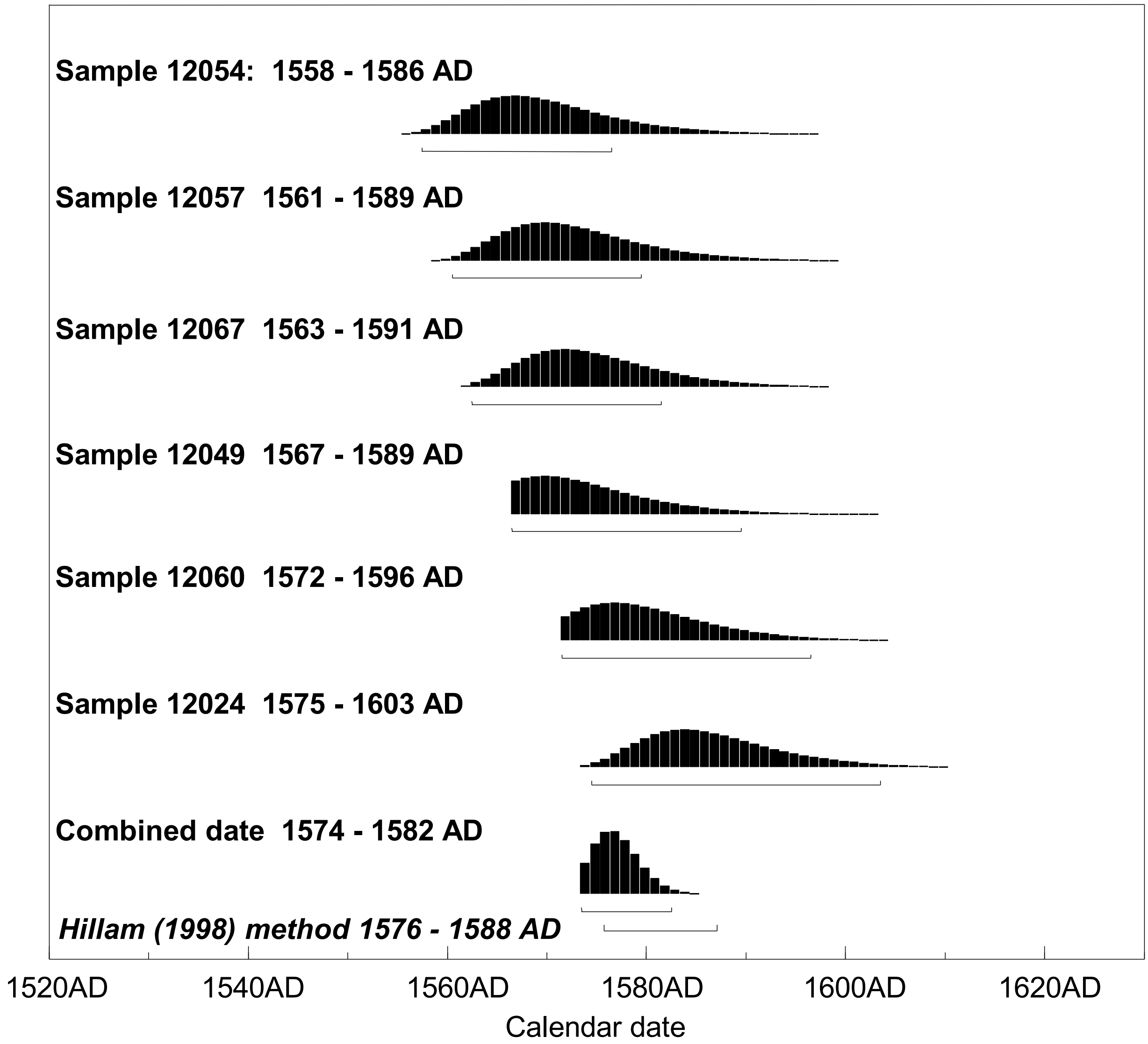

Figure 3: Probability distributions for the felling date of timbers from the Great Kitchen, Windsor Castle.